

# Exhibit D

## Exhibit D

DATE SUBMITTED: [REDACTED]

SUGGESTED TITLE FOR THE INVENTION: Phased Array Sound System

**BRIEF DESCRIPTION OF THE INVENTION:** A sound reproduction method that provides a means of controlling sound intensity (volume) and program content at arbitrary locations within a given listening space.

### DETAILED DESCRIPTION OF THE INVENTION:

#### Overview:

A sound reproduction method employing multiple loud speakers (or ultrasonic transducers as described in an alternate embodiment) and digital signal processing technology to produce phased array sound fields where the intensity of the sound at any location in the defined listening space can be controlled by constructive superposition of the sound waves emanating from the loud speaker or transducer array. Constructive superposition of the sound waves is achieved by using a digital signal processor (DSP), or other suitably fast embedded processor system, to insert variable delays into the input signal going into specific speakers in the array. The delays are calculated using the distance of each speaker or transducer to a specific target location within the listening space. Using the distance to the target and the assumed average speed of sound at room temperature and room pressure, the maximum time of flight for the sound signal of the farthest speaker from the target in the array is calculated, then the exact amount of additional delay (within given tolerances) is inserted into the input signal for each closer speaker in the array to the target. This method achieves the result that the wave front from each speaker arrives at the target at the same time and roughly in phase. Due to superposition the amplitudes of wavefronts will algebraically add. Since sound intensity or volume is a function of the square of the signal amplitude, very significant sound intensities at the target can be achieved for reasonable sized arrays. An array of 10x10 loud speakers has the theoretical limit of achieving an increase of intensity at the target of 10000 times or 40dB over the volume of any one speaker as heard from the same location. The sound intensity at locations other than the target location is a function of sound addition that is not in phase but randomly distributed in phase and timing and is proportional to the number of speakers or transducers in the array rather than the square of the number of speakers or transducers. The resulting difference in sound intensity levels that might be achieved, if for instance each speaker were to be operated at a sound level of 0dB or the threshold of audibility would be that at everywhere in the listening space except the target, the out-of-phase and unintelligible signal would have a sound volume of a faint whisper or the sound of rustling leaves while at the target a fully intelligible, focused and in-phase signal would have a sound intensity level nearly equal to that of normal conversation. The target within the listening area can be thought of as focal point for a parabolic or elliptical mirror. As proof of this concept, a mechanical analog of this method exists in the form of a whispering gallery. The whispering gallery is an elliptical shaped room, and as such the room has two focal points. Sound emanating from one focal point is bounced around the room and is reconstructed in phase and amplitude at the other focus. Thus, a seemingly private conversation initiated unwittingly at one focus, that is essentially inaudible at most every where else in the whispering gallery, is reconstructed and is perfectly audible at the other focus some distance away. The advantage of the electronic embodiment of this concept is that there can be multiple focal points or targets, target locations can be created or eliminated dynamically, the volume or sound intensity at each target can be made essentially independent of the other targets and the program content (what is heard) at each target can be made essentially independent of the other targets. Furthermore, the targets can be moved dynamically within the listening space and with some additional effort can be made to track with independent listeners within the room.

#### Theory of operation:

The first type of system discussed will be referred to as a base-band system. In the base-band system it is the actual audio program that is focused into a specific target location. (Please refer to Figure 1) This is the block diagram of the simplest embodiment of the process which is a base-band, single target focus, single audio program.

#### A to D:

file c:\winword\invrec1.doc

1

Print date: [REDACTED]

The analog audio program source is digitized by the A to D block in Figure 1. The sampling need only satisfy the Nyquist criteria for the audio spectrum audible to human beings. Since the top end of human hearing is 20 KHz, the Nyquist criteria would dictate a sampling rate of 40 KHz. There is no need to over-sample as the signal is only subjected to a simple delay algorithm and is low pass filtered in the analog domain. This embodiment could use a 44 KHz sampling rate. This is a common Pulse Code Modulation (PCM) [".WAV"] file format sampling rate used in readily available sound card technology.

#### PCM Wave Form Buffer:

This buffer receives a constant stream of data. Each piece of data is a Pulse Code Modulated (PCM) wave form sample. Each sample is roughly 22 uSec newer than the previous sample at the above mentioned sampling rate. The buffer is a simple shift register, such that with the arrival of each new PCM wave form sample, the previous sample is shifted by an offset of one into the buffer. The net effect is to produce a delay line of PCM waveform samples whereby the older the sample is in time, the larger the offset is into the buffer. The buffer is the means by which delay is entered into wave form signal that is ultimately applied to the output speakers. The size of the buffer need only be as large as the number of samples that would satisfy the following equation:

$$\text{Buffer size (in samples)} = R_s \times (d_m/V_l)$$

Where:

$d_m$  is the maximum distance that sound may have to travel for a given listening room size in meters (m).

$V_l$  is the speed of sound in air in meters per second (m/s).

$R_s$  is the sample rate of the PCM wave form in samples per second (sample/sec)

For an average 12' by 12' room and a common sampling rate of 44 KHz the size of the buffer works out to be 511 samples long. 512 is a nice binary size that would be used in the preferred embodiment.

#### The Data Pointer Offset Array:

There is a one to one correspondence between each element of the Data Pointer Offset Array and a speaker in the NxN Speaker Array of Figure 1. The correspondence is such that each element of the Data Pointer Offset Array holds a pointer to a specific address in the PCM waveform buffer for a given speaker in the NxN Speaker Array. In other words the  $k^{th}$  element of the Data Pointer Offset Array will be uniquely associated with the speaker  $S_{ij}$ , the speaker in the  $i^{th}$  row and  $j^{th}$  column of the NxN speaker array. The specific address held in the Data Pointer Offset Array represents an offset into the PCM waveform buffer. This offset represents the specific time delay that is introduced in the audio program signal that is sent to that specific speaker. Any given data element of the pointer offset array will hold an offset into the PCM waveform buffer that is given by the following formula:

$$\text{Offset} = R_s \times T_{adj}$$

Where:

$R_s$  is the sample rate of the PCM wave form in samples per second (sample/sec)

$T_{adj}$  is the additional time delay (in seconds) necessary at speaker location  $S_{ij}$ .

It is given by the formula:

$$T_{adj} = T_{mc} - T_{slj}$$

Where:

$T_{mc}$  is the time delay due to the time of flight at the speed of sound over the distance between the target and the furthest speaker from the target.

$T_{slj}$  is the time delay due to the time of flight at the speed of sound over the distance between the target and the speaker  $S_{ij}$ .

#### The Address Generator:

The address generator is actually a function of the program that runs on the CPU. It is shown on the block diagram for purposes of clarity. The address refers to the address that is sent to the analog

multiplexor (Analog MUX in Figure 1) and is used to route the output of the digital to analog converter (DAC in Figure 1) to the correct speaker in the NxN speaker array.

#### The CPU:

The CPU is an embedded processor of some variety that is suitably fast to perform the function listed below. In the preferred embodiment a Digital Signal Processor (DSP) would be used.

The primary loop of the operating program performs the following operations:

- 1.) Generates an address that selects a specific speaker  $S_{ij}$  in the NxN speaker array. The address is generated by means of a simple incrementing process that starts from zero and is incremented by one at the completion of each loop in the primary loop of operation.
- 2.) Uses the address generated in step one as an offset address into the Data Pointer Offset array to indirectly address and read the specific waveform sample held in the PCM waveform buffer into the accumulator of the CPU. At this point, the data sample can be scaled up or down for volume control.
- 3.) Outputs the address generated in step one to the output port that is connected to the Analog MUX.
- 4.) Outputs the data now held in the accumulator to the output port that is connected to the DAC.
- 5.) Increments the address from step one by one and loop back to step one.

The entire loop must be repeated for every speaker and all speakers must receive an output sample from the DAC within the waveform sampling period or 22  $\mu$ s for the preferred embodiment.

Use of a digital signal processor can be helpful in combining steps such as the incrementing step so as to save computational time. Depending on the size of the speaker array, multiple CPUs may be used to maintain computational throughput. Fortunately, since the steps involved in the primary loop are fairly rudimentary the processors needed can be of an inexpensive type.

The pointers held in the Data Pointer Offset array are calculated based upon the known location of the target in relation to the speaker array. If the target focus location is fixed, the data pointers in this array need only be calculated once. If the target focus location is moved, the pointers in this array need to be updated. If the target is to be tracked while it is in motion, the pointers in this array must be updated continuously.

#### The speaker array:

The speaker array in the preferred embodiment consists of an NxN array of small (2 to 3 inch diameter) high quality audio speakers that are positioned in a rectilinear grid pattern in the ceiling of the subject listening room. The speakers need to be selected for their ability to be non-resonant, that is a frequency response that is reasonably flat in the passband of interest (in this case the passband of human hearing) and non-directional so as to provide a hemispheric wave front or a reasonable facsimile thereof.

(Please refer to Figure 2) This is the block diagram of a slightly more elaborate embodiment of the process which is a base-band, multiple target focus, single audio program.

In this situation, everything is the same as for the simplest embodiment except there are multiple Data Pointer Offset Arrays. There is one array for each target focus. Each array holds a set of pointers that produce the specific delays that will focus the sound at a predetermined location in the listening area. In this case the CPU must indirectly address and read a PCM wave form sample through each of the multiple Data Pointer Offset Arrays for each speaker location. The multiple samples are summed and output to the DAC. Multiple targets are created by a simple process of superposition of the waveform samples on the array of speakers.

(Please refer to Figure 3) This is the block diagram of a still more elaborate embodiment of the process which is a base-band, multiple target focus, multiple audio program.

In this case everything is the same as the system pictured in Figure 2 except there are multiple PCM Wave form buffers. The waveform samples for a given target location are indirectly accessed from the waveform buffer that holds the desired audio program. This is a simple matter of setting a different base address when performing the indirect address through the data pointer offset array.

file c:\winword\invrec1.doc

3

Print date: [REDACTED]

(Please refer to Figure 4) This is the block diagram of a complete embodiment of the process which is a base band, multiple target focus, multiple audio program with listener tracking.

The tracking system that is described here is by no means the only tracking system that could be used. Other listener tracking system that might be used are simply alternate embodiments of the same system.

Here a tracking system is implemented that creates an interrogating target focus. The interrogating target focus is a sub-audible tone pattern that is localized in three dimensional space and continuously scanned through the listening room. Listeners would be required to wear a wireless microphone that was omnidirectional and sensitive to the frequencies in the tone pattern. A listener's location within the listening room is fixed in relation to the speaker array by sensing the time at which the signal from a given microphone reaches its maximum. The location of the interrogating target focus is known by the contents of the data pointer offset array for the interrogating tone. At the time when the maximum is sensed, a copy of the data pointer offset array for the interrogating tone is placed in the array for the listener target.

It should be noted that this method reduces the computational load of the CPU since it eliminates the need to calculate the delays to be programmed into the Data Pointer Offset Array as the target locations are identified empirically as the interrogating tone is scanned through the room. The process of scanning a tone target is a simple matter of incrementing in a predetermined fashion the data elements of the Data Pointer Offset Array for the target focus to be scanned through the room.

The disadvantage of this method is that the listeners are still required to wear some kind of appliance (i.e. wireless microphone). It would be best if the items are worn high on body, the shoulder or collar would be optimal. In light of current technology and the power requirements for wireless transmission, it is conceivable that the item could be made to look like a small pin or jewelry.

#### Alternate embodiment

The base band system has some minor drawbacks. One of these is that the reconstruction of lower wavelengths occurs over a larger volume of physical space than higher wavelengths so that the area of increased sound pressure may overlap into another target area, if the target spacing is not held to specific limits. Another is that the dynamic range of the audio program may be limited if the goal of relative inaudibility is sought for non-target locations. Even though the signal is likely to be unintelligible, the sound intensity at frequencies in the audible range is proportional to the number of speakers in the array and the aggregate power output level of the speaker array. These issues can be minimized to a great degree through implementation of the alternate embodiment of the phased array sound system. In this embodiment, the audio program is directed to the target location via a phase modulated ultrasonic carrier wave. The modulated carrier is of the form " $A \cos(\omega_c \tau + \theta)$ " where  $A$  is the amplitude,  $\omega_c$  is the frequency  $2\pi f_c$ , and  $\theta$  is varied between  $-\pi/2$  and  $+\pi/2$  (in radians). The signal that modulates the carrier wave is the audio program signal that is desired to be heard at the target location. The audio program signal is then demodulated at the target location in free air through the process of heterodyning with an unmodulated carrier at the same frequency and zero phase displacement. This is achieved by placing, at the target location, a pure sinusoidal wave of the form  $A \cos(\omega_c \tau)$ . It is important that the frequency of both sinusoidal signals is exactly the same and that the unmodulated wave form has 0 phase offset in relation to the modulated waveform when  $\theta$  in the modulated wave form is 0. This heterodyning results in a translation from phase modulation to amplitude modulation. The size of the respective signals are controlled to ensure that the amplitude modulation is maintained at or below 100%. The original base band audio program is recovered through the use of square law effect amplification to introduce a non-linearity in the amplitude modulated signal. The non-linearity is achieved through the process of transducing the 40kHz electrical signal to ultrasound. During this process, square law amplifiers are used to introduce a small non-linearity such that the sinusoids are not perfectly symmetrical. The resulting asymmetry will produce a wave form for the modulated (as well as the unmodulated) waveform that has positive (negative) excursions that are slightly larger than the negative (positive) ones (see Figure 5). Once these conditions are met at the target location, the result is an average signal that follows the envelope of the modulated signal. Since the amplitude modulated signal is on a carrier wave that is well

file c:\winword\invrec1.doc

above the audible range, the human ear will provide what amounts to a low pass filtering of the average envelope of the amplitude modulated signal. The sound perceived at the target location is that of the original base-band audio program.

Greater dynamic range in the reconstruction of the signal is possible due to the fact that sound pressure in the non-target locations has the vast majority of its signal energy at frequencies well above the range of human hearing. The base-band signal is reconstructed in the very localized volume of the target focus location. If the carrier wave is 40 KHz, the target focus is a spherical volume of roughly .33 inches in diameter. Thus, target locations can be positioned much more closely with no overlap. The air in this volume is compressed and rarefied at all the signal frequencies contained in the amplitude modulated signal. Air outside this focus location receives energy from the focus in much the same way as a speaker imparts sound energy to the air in its immediate vicinity. The ultrasonic transducers that produce the sound can be run at very high signal output intensities. (i.e. 160dB) Even though it is unlikely that such high sound output powers would actually be used, it is an indication that significant dynamic range output power could be achieved.

A course implementation of this scheme is pictured in Figures 6 and 7. Figure 7 shows a system very similar to that pictured in Figure 4 with the exception that the speaker array has been replaced with an array of ultrasonic transducers. Communication to each element of the array is now done via a digital serial link and there is a high accuracy radio frequency synchronizing pulse that is used to ensure accurate phase alignment for all of the transducer elements of the array. Figure 6 shows a typical array element which has three major components. The components are an application specific integrated circuit, a square law amplifier and an ultrasonic transducer. The application specific IC produces the phase modulated and/or unmodulated electrical signal that will have the proper phase relationship to the other transducers in the array so as to produce the desired aiming of the signal into the listening space. It is comprised of: 1) a controller that receives digital data from a high speed serial port that is addressed to this specific transducer location. 2) Single element data buffers that hold pointer offsets into the carrier wave buffer. These offsets are updated by the data received from the serial port. 3) A time base that is phase locked to the synchronizing signal. 4) A waveform buffer that contains pulse code modulated waveform samples that represents one full cycle of the carrier wave sinusoid. 5) A summing function that produces a net pointer into the carrier wave sinusoid buffer and 6) a summing function that sums the amplitude of the carrier wave sample for each target. 7) a digital-to-analog converter. The square law amplifier provides the bias voltage and driving voltage for the ultrasonic transducer. It also introduces a small non-linearity into the signal such that the carrier wave is slightly asymmetrical. It should be noted that the non-linearity may be encoded into the PCM waveform samples of the carrier wave sinusoid thus eliminating the need for a special non-linear amplifier but the square law amplifier is shown here for clarity. An amplifier is needed in any case to provide the special bias and driving voltages for the transducer. The ultrasonic transducer is a commercially available (Polaroid) electrostatic transducer. In this case a 7000 series. It is operated near its peak efficiency of 55KHz. The aperture plate is employed to act as a beam spreader. Without the aperture plate the beam pattern from the transducer is highly directional. A 50% decrease of the aperture produces a doubling of the beam angle. The net effect of the aperture (once properly sized) is to act like a pin hole lens and produce a hemispherical wave front to as great a degree as possible.

The operation of the transducer array element is as follows: in the base-band system the audio program signal had actual delays introduced for specific array speakers to produce convergence of the signal at a specific point in the listening room. In order for the carrier wave to converge at a specific target location, a similar concept is employed, however, for a continuous sinusoid this delay reduces down to a whole number of cycles and a residual phase. It is only necessary to modify the phase of the output sinusoid for each element of the array in order to produce constructive superposition at the desired target location for the unmodulated sinusoid. The time base is a phase-locked loop, locked to the time base synchronizing signal. It is used to increment a pointer through the carrier wave sinusoid PCM buffer. All elements of the array increment to the next PCM sample in lock with the synchronizing signal. The speed of incrementing is such that the entire buffer is strobed through at a rate of 55KHz. The residual phase necessary for this particular element is read from the target phase offset buffer. The amount of phase angle offset is in the form of a signed integer number of samples. Thus, the number read from the "Target/channel modulation phase offset single data element buffer" is summed with the pointer

file c:\winword\invrec1.doc

produced by the time base and is used to retrieve the PCM sample from the carrier wave sinusoid buffer. Audio program phase modulation of the carrier wave is incorporated into the target phase offset with the phase modulation offset presented to the summing function delayed in time by an amount dictated by the central CPU of Figure 7. Multiple targets and multiple channels are for each element achieved by a simple process of summing the samples that are mapped to this particular transducer element of the array.

The phase output of each element of the array, when considered across the entire array, can be thought of as a time varying two dimensional Fourier Transform representing the spatial frequencies of the target locations. The array is, in essence, a phase-encoded hologram where the reference/playback beam is the ultrasonic carrier wave.

Figure 7 provides the target-location phase offset calculation and audio program array element delays for each element of the array. The computational load is only slightly larger for the CPU in this system as compared to the baseband system since the delay offset pointer for each array element must be converted into a signed integer number of carrier wave samples before being transmitted to the particular array element that is mapped to that pointer. Since the targets in question don't move with great rapidity in relation to the computational reaction time of the CPU, this relationship of carrier wave offset samples to data pointer offsets (as pictured in Figure 7) is treated as a lookup table. In this case, the data pointer offset array points to a data element that is the sum the PCM waveform buffer sample (that has been converted to a carrier wave phase offset number) and the target phase offset (that has also been converted to a carrier wave offset number). An overall communications efficiency is realized by allowing the summing to occur at the central CPU so that only one number per channel per ASIC needs to be transmitted to each transducer in the array. In other words if 3 active targets/channels are being transmitted to in the listening room then the central CPU must update each ASIC at each transducer location in the array with 3 offset numbers.

Listener tracking is done in an identical fashion to the base-band system.

#### **KNOWN PRIOR PUBLICATION:**

None relating to the production of "real image" of sound.

While there is very active work in the area of "3D" sound; these efforts are all aimed at exploiting the psycho-acoustic effects of time and phase relationships produced by stereo and Dolby type systems. These systems take advantage of the biophysical phenomena Interaural Time Differences (ITD), Interaural Intensity Differences (IID) as well as Head Related Transfer function Function (HRTF) to create "virtual images" of sound.

#### **ADVANTAGES OVER OTHER PRIOR ART:**

Current and prior art related to public address as well as other sound reproduction systems has concentrated on creating a homogenous soundfield even when specific elements of an audio program were to have an identifiable location in the soundfield, it was intended to be heard by all in the same listening room. Aside from the use of headphones, private listening was always a matter of keeping the volume down. Prior art has never made available distinct and controllable volume levels for individuals in the same listening room. Aside from the use of headphones, prior art has never been able to provide different audio programs to the inhabitants of the same listening room. As examples of the usefulness of this process, consider multilingual school rooms or auditoriums where listeners if properly equipped (using a system with listener tracking) or seated in the proper location (in a non-tracking type system), might hear the presenter in his or her own language without the use of cumbersome headphones. In the entertainment realm, it might be possible to have both edited and non-edited versions of motion picture film dialog presented to the same audience at the same time. Entirely different plot lines may be able to be presented to different portions of the audience. In public places, one juke box might be able to play several different songs at the same time. Hands-free phone operation might be achieved in open office environments while still maintaining private conversation. Buildings so equipped could take advantage of listener tracking to automatically route telephone and intercom signals to the desired recipient without the need of a handset. The conversation could be full duplex if the tracking system were of the wireless microphone type.

file c:\winword\invrec1.doc



Since the sound produced is a "real image" (it actually comes from the location it appears to come from) there are many opportunities for sound reproduction and special effects of video games, multimedia presentations and high fidelity music. Specifically, prior art did not have the ability to place the "real sound image" of an audio source in three dimensional space.

**ASSERTION OF INVENTORSHIP:** I conceived and refined the above stated concepts and ideas by myself over the period of time from [REDACTED]

**CONCEPTION DATE:** [REDACTED]

**DATE FIRST SKETCH OR DRAWING WAS MADE:** [REDACTED]

**DATE FIRST DISCLOSED TO ANOTHER:** [REDACTED]

I have read and do understand this document consisting of 7 pages plus 6 figures.

Daniel R. Butler

Daniel R. Butler

State of Wisconsin  
County of Waukesha

Signed or attested before me on [REDACTED] by Daniel R. Butler

Monica J. Cody

My commission expires [REDACTED]

file:winword\invrec1.doc

Print date [REDACTED]



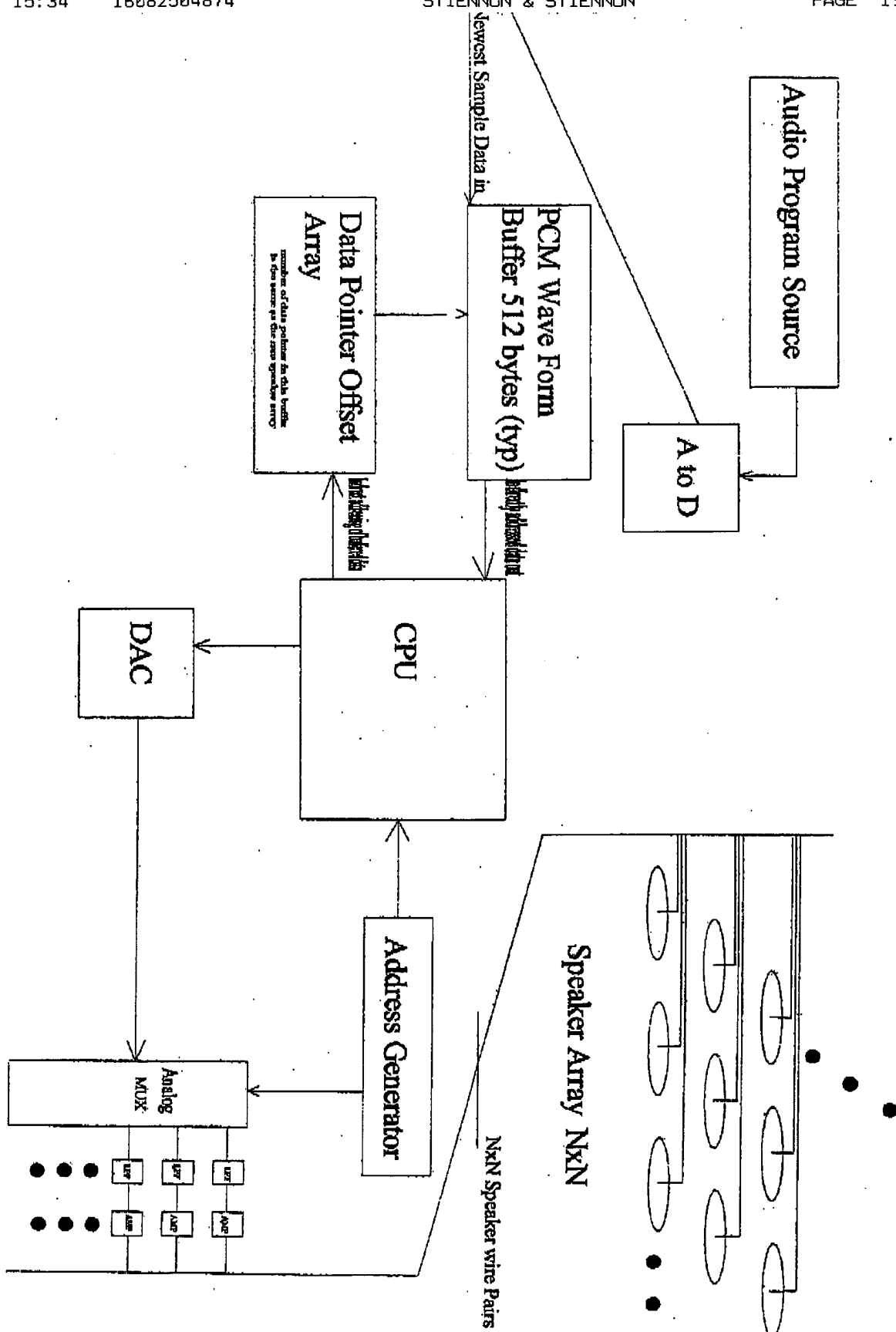


Figure 1.

Phased Array Sound System Block Diagram

(Single Target System) J.P.M

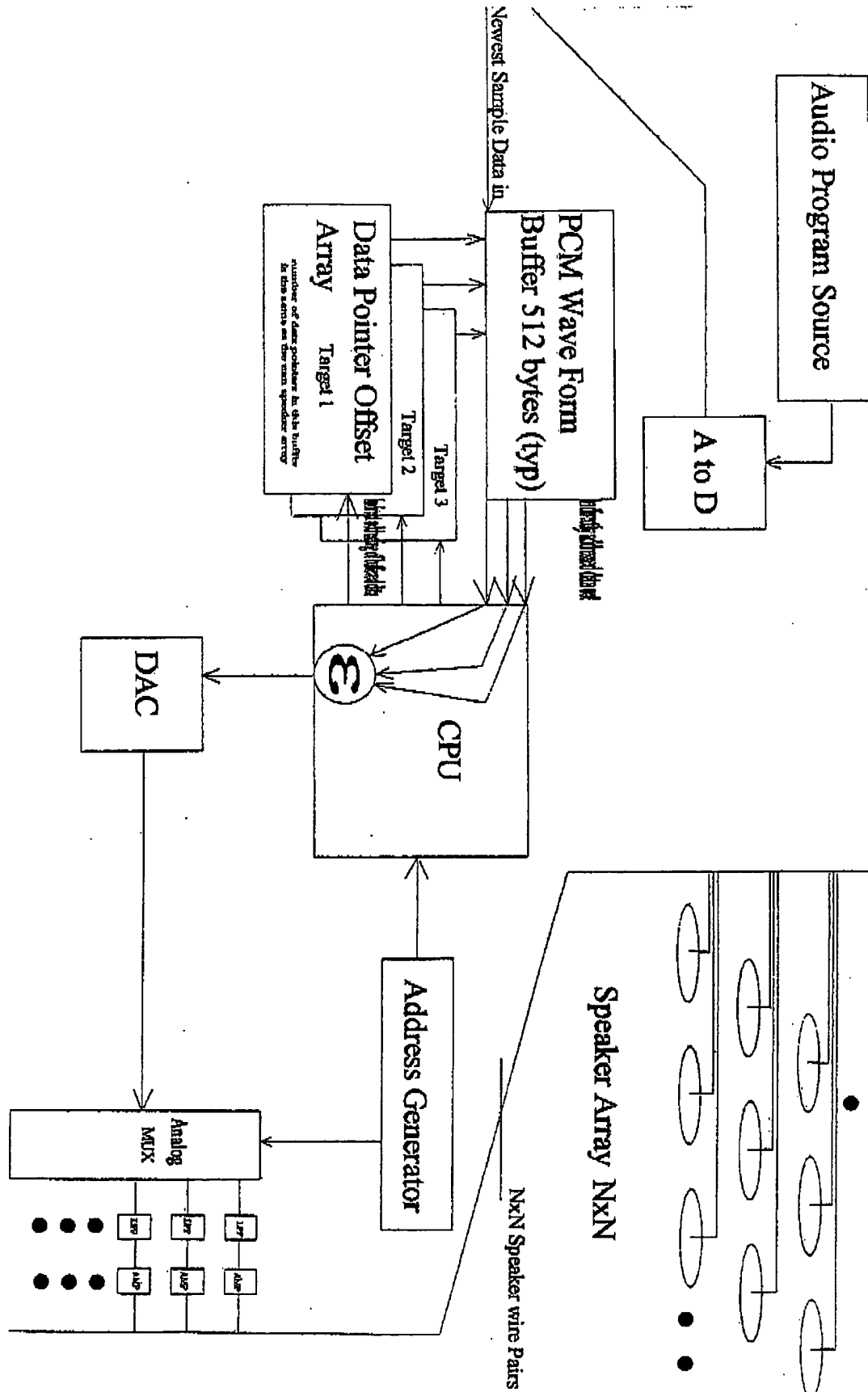


Figure 2.

Phased Array Sound System Block Diagram (Multiple Target System) J.P.M

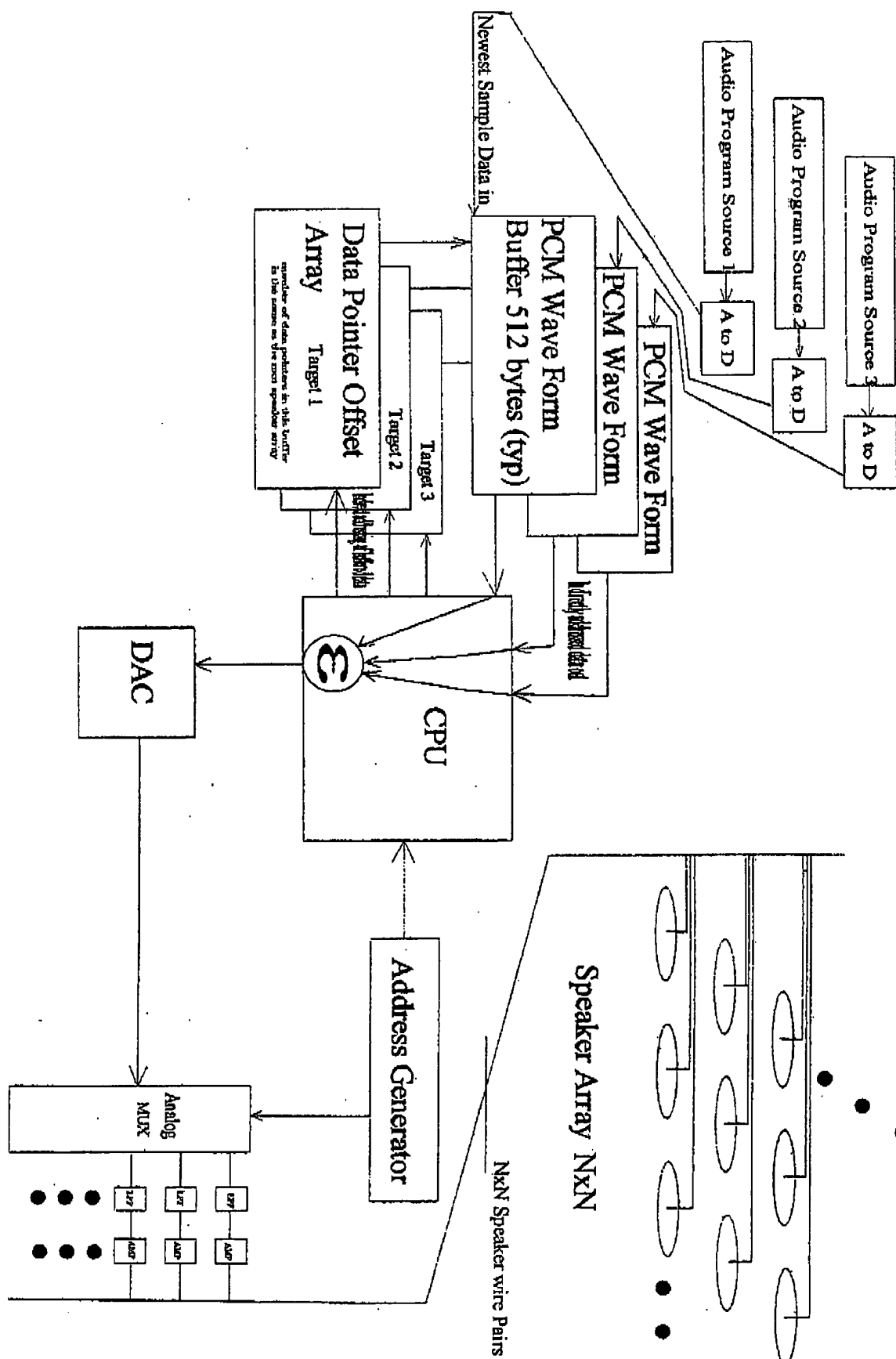


Figure 3.

Phased Array Sound System Block Diagram (Multiple Target Multiple Program System)

J.P.M.

NxN Analog Channels

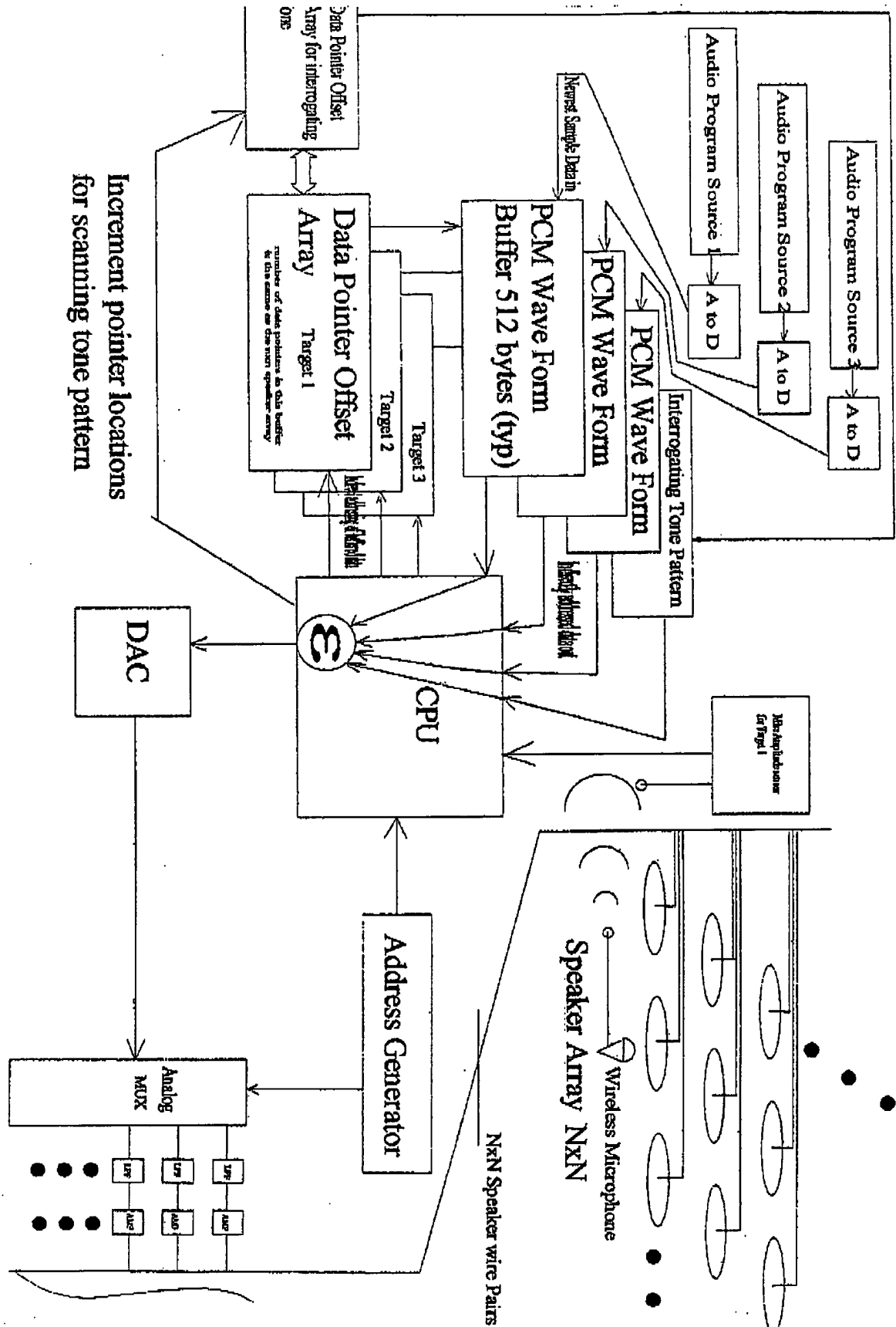


Figure 4.

Phased Array Sound System Block Diagram

(Multiple Target Multiple Program System with listener tracking)

J.P.M

NxN Analog Channels

$$\text{Input} = \sin(2\pi f t) \cos(2\pi f t)$$

$$\text{output} = (\sin(2\pi f t) \cos(2\pi f t))^2$$

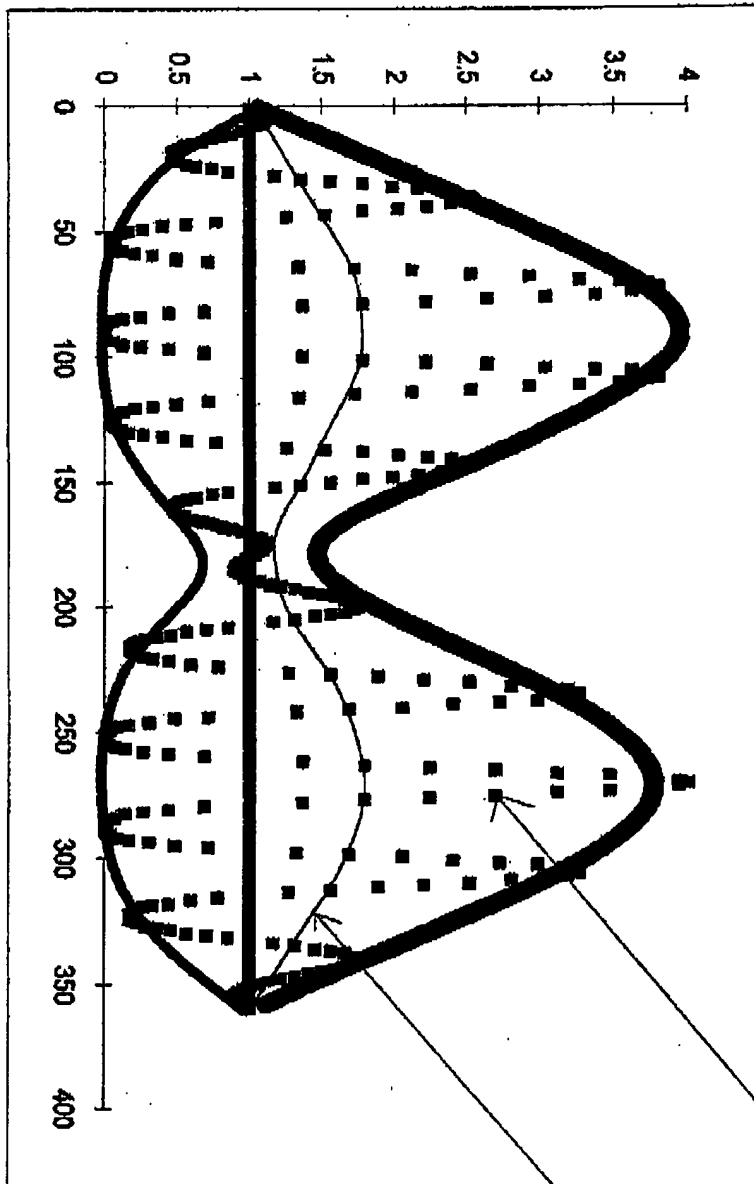


Figure 5

Illustration of the operation of square law demodulation. The output is the net envelope averaged over many cycles

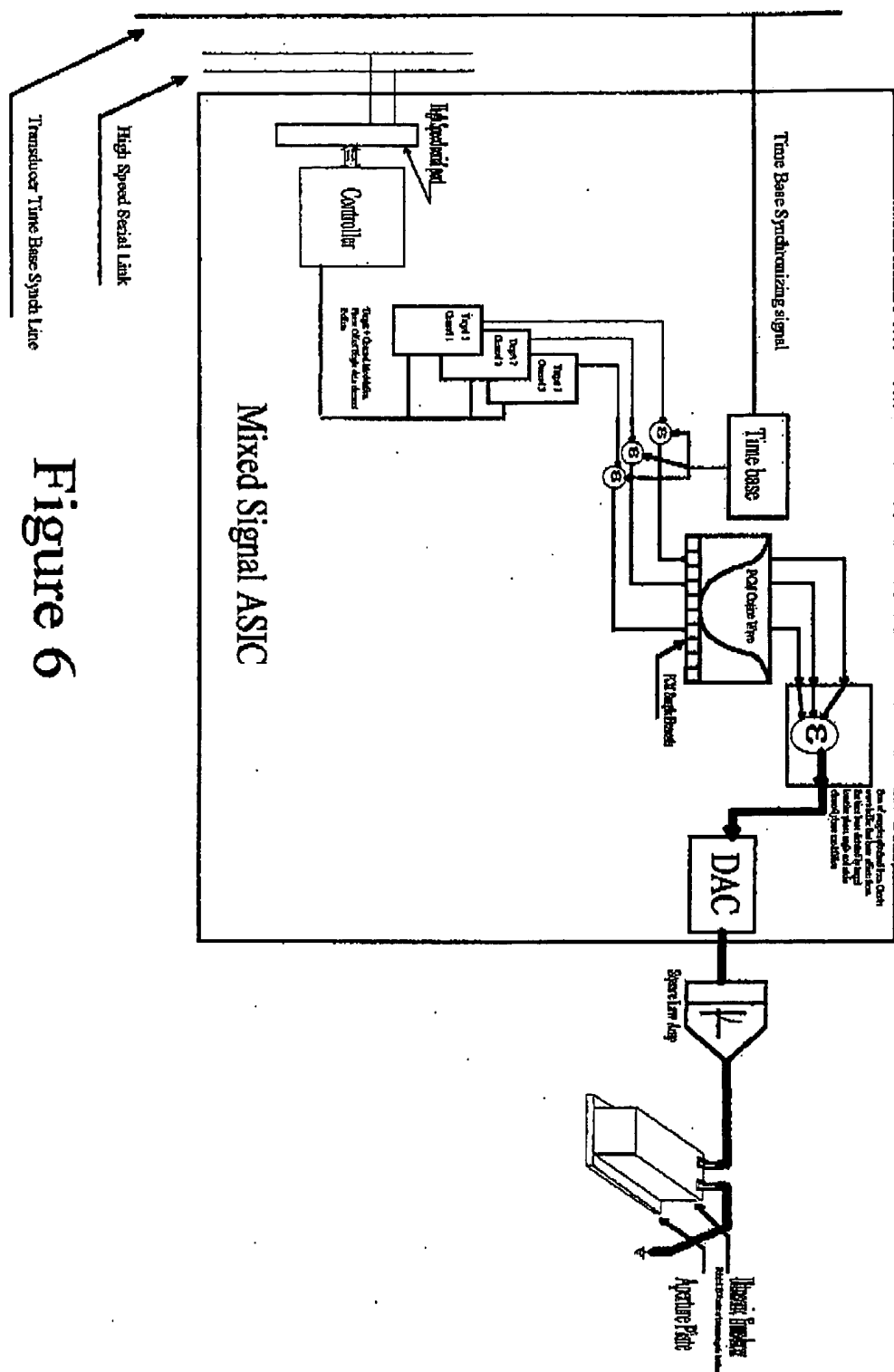


Figure 6

Typical Array Element Composed of an ASIC, Non Linear Amp and Ultrasonic Transducer



**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**